

# A DIMENSION GAP FOR CONTINUED FRACTIONS WITH INDEPENDENT DIGITS - THE NON STATIONARY CASE

ARIEL RAPAPORT

**ABSTRACT.** We show there exists a constant  $0 < c_0 < 1$  such that the dimension of every measure on  $[0, 1]$ , which makes the digits in the continued fraction expansion independent, is at most  $1 - c_0$ . This extends a result of Kifer, Peres and Weiss from 2001, which established this under the additional assumption of stationarity. For  $k \geq 1$  we prove an analogues statement for measures under which the digits form a  $*$ -mixing  $k$ -step Markov chain. This is also generalized to the case of  $f$ -expansions. In addition, we construct for each  $k$  a measure, which makes the continued fraction digits a stationary and  $*$ -mixing  $k$ -step Markov chain, with dimension at least  $1 - 2^{3-k}$ .

## 1. INTRODUCTION

Let  $X$  denote the set of irrational numbers in  $(0, 1)$ . It is well known each  $x \in X$  has a unique continued fraction expansion of the form

$$x = \frac{1}{A_1(x) + \frac{1}{A_2(x) + \frac{1}{A_3(x) + \dots}}},$$

where  $A_1(x), A_2(x), \dots$  are positive integers. Given a probability measure  $\nu$  on  $X$ , each  $A_n$  defines a random variable on  $(X, \nu)$  and the digits  $\{A_n\}_{n=1}^\infty$  form a discrete time stochastic process.

In 1966, Chatterji [Ch] has shown every probability measure  $\nu$  on  $[0, 1]$ , which makes the digits in the continued fraction expansion independent variables, is singular with respect to the Lebesgue measure. In 2001, Kifer, Peres and Weiss [KPW] have proven that  $\dim_H \nu \leq 1 - c$ , if in addition the digits are identically distributed. Here  $0 < c < 1$  is a global constant, independent of  $\nu$ , and  $\dim_H \nu$  denotes the Hausdorff dimension of  $\nu$ , which is defined in Section 2 below. In this paper we show the result from [KPW] remains true, even if the digits are independent but not necessarily identically distributed.

---

*Date:* March 9, 2017.

*2000 Mathematics Subject Classification.* Primary: 11K55, Secondary: 37C45.

*Key words and phrases.* Continued fractions, Hausdorff dimension.

Supported by ERC grant 306494.

Assuming  $A_1, A_2, \dots$  are i.i.d. with  $\mathbb{E}[\log A_1] < \infty$  and  $H(A_1) < \infty$ , where  $H(A_1)$  is the entropy of  $A_1$ , Kinney and Pitcher [KP] have proven that

$$(1.1) \quad \dim_H \nu = \frac{H(A_1)}{-\int_0^1 \log x^2 d\nu}.$$

The Gauss measure

$$\mu_G(E) = \frac{1}{\log 2} \int_E \frac{dx}{1+x}$$

is the unique equilibrium state of the Gauss map  $Tx = \frac{1}{x} \pmod{1}$  with respect to the function  $x \rightarrow \log x^2$ . This follows from the thermodynamic formalism approach of Walters [Wa1]. Hence under the i.i.d. assumption

$$0 = \int_0^1 \log x^2 d\mu_G(x) + h_{\mu_G}(T) > \int_0^1 \log x^2 d\nu(x) + h_\nu(T),$$

where  $h_\eta(T)$  is the entropy of  $T$  with respect to a  $T$ -invariant measure  $\eta$ . Since  $h_\nu(T) = H(A_1)$ , we get from (1.1) that  $\dim_H \nu < 1$  in this case. When  $A_1, A_2, \dots$  are not identically distributed the formula (1.1) is no longer valid, and so it is not even clear that  $\dim_H \nu$  is strictly less than 1. As mentioned above, we shall show that there exists a global constant  $c_0 > 0$  such that  $\dim_H \nu \leq 1 - c_0$ , assuming  $A_1, A_2, \dots$  are independent.

We actually prove more generally that for every integer  $k \geq 0$  there exists  $0 < c_k < 1$ , which depends only on  $k$ , such that  $\dim_H \nu \leq 1 - c_k$  if the digits form a  $k$ -step Markov chain which is  $*$ -mixing. This is the main result of this paper. The  $*$ -mixing condition was introduced in [BHK], and is a bit less restrictive than the more familiar  $\psi$ -mixing condition. The definitions are given in Section 2. In the last section we generalize our main result to the case of  $f$ -expansions.

Given  $k \geq 0$  it was shown in [KPW] that there exists  $0 < c'_k < 1$ , for which  $\dim_H \nu \leq 1 - c'_k$  whenever  $\nu$  makes the digits a stationary and ergodic  $k$ -step Markov chain. Our proof is a modification of the argument given there for this result. We shall also construct for each  $k$  a measure  $\nu_k$ , under which the digits form a stationary and  $\psi$ -mixing  $k$ -step Markov chain, with  $\dim_H \nu_k \geq 1 - 2^{3-k}$ . This of course shows  $c_k$  and  $c'_k$  are at most  $2^{3-k}$ .

The rest of the paper is organized as follows. In Section 2 we give some necessary definitions and state our results. In Section 3 we establish a uniform bound on the dimension of subsets of  $X$ , which are defined via certain digit frequencies. This is the key ingredient in the proof of our main result, which is carried out in Section 4. In Section 5 we construct the measures  $\nu_k$  mentioned above. In Section 6 we generalize our main result to the setup of  $f$ -expansions.

**Acknowledgement.** This paper is a part of the author's PhD thesis conducted at the Hebrew University of Jerusalem. I would like to thank my advisor Professor

Yuri Kifer, for suggesting to me the problem studied in this paper, and for many helpful discussions.

## 2. PRELIMINARIES AND RESULTS

First, we define the mixing conditions mentioned above. Given random variables  $\{A_i\}_{i \in I}$ , all defined on the same probability space, denote by  $\sigma\{A_i\}_{i \in I}$  the smallest  $\sigma$ -algebra with respect to which each  $A_i$  is measurable.

**Definition 2.1.** A sequence of random variables  $\{A_n\}_{n=1}^\infty$  is called  $*$ -mixing if there exist an integer  $N \geq 1$  and a real valued function  $f$ , defined on the integers  $n \geq N$ , such that

- $f$  is non-increasing with  $\lim_{n \rightarrow \infty} f(n) = 0$ , and
- if  $n \geq N$ ,  $m \geq 1$ ,  $E \in \sigma\{A_1, \dots, A_m\}$  and  $F \in \sigma\{A_{m+n}, A_{m+n+1}, \dots\}$  then

$$|\mathbb{P}(E \cap F) - \mathbb{P}(E)\mathbb{P}(F)| \leq f(n)\mathbb{P}(E)\mathbb{P}(F).$$

If such an  $f$  exists for  $N = 1$  the sequence is said to be  $\psi$ -mixing.

*Remark 2.2.* A sequence of independent random variables is clearly  $\psi$ -mixing. It is not hard to show that the  $\psi$ -mixing condition is satisfied for a finite state Markov chain  $\{A_n\}_{n=1}^\infty$ , with state space  $S$ , for which

$$\inf\{\mathbb{P}(A_{n+1} = j \mid A_n = i) : n \geq 1 \text{ and } i, j \in S\} > 0.$$

Examples of  $*$ -mixing countable state Markov chains can be found in Section 3 of [BHK]. Another important example of a  $\psi$ -mixing sequence is obtained by the continued fraction digits with respect to the Gauss measure  $\mu_G$  (see [Ad] or [He]).

Set  $X = (0, 1) \setminus \mathbb{Q}$  and for each  $x \in X$  and  $i \geq 1$  let  $\alpha_i(x) \in \mathbb{N} := \{1, 2, \dots\}$  be the  $i$ 'th digit in the continued fraction expansion of  $x$ , i.e.

$$x = \frac{1}{\alpha_1(x) + \frac{1}{\alpha_2(x) + \frac{1}{\alpha_3(x) + \dots}}}.$$

Given  $a_1, a_2, \dots \in \mathbb{N}$  denote by  $[a_1, a_2, \dots]$  the unique  $x \in X$  with  $\alpha_i(x) = a_i$  for  $i \geq 1$ . For  $E \subset X$  write  $\dim_H(E)$  for the Hausdorff dimension of  $E$ . Given a Borel probability measure  $\nu$  on  $X$  its Hausdorff dimension is defined by

$$\dim_H(\nu) = \inf\{\dim_H(E) : E \subset X \text{ is a Borel set with } \nu(E) = 1\}.$$

The following theorem is our main result.

**Theorem 2.3.** Let  $\{A_n\}_{n=1}^\infty$  be  $\mathbb{N}$ -valued random variables and let  $k \geq 0$ . Assume  $\{A_n\}_{n=1}^\infty$  is a  $k$ -step Markov chain (when  $k = 0$  this means  $A_1, A_2, \dots$  are independent) which is  $*$ -mixing. Let  $\nu$  be the distribution of the random variable  $[A_1, A_2, \dots]$ . Then  $\dim_H(\nu) \leq 1 - c_k$ , where  $0 < c_k < 1$  is a constant depending only on  $k$ .

*Remark 2.4.* As mentioned in the introduction, it was shown in [KPW] that there exists  $0 < c'_k < 1$ , for which  $\dim_H \nu \leq 1 - c'_k$  whenever  $\nu$  makes the continued fraction digits a stationary and ergodic  $k$ -step Markov chain.

It might be desirable to estimate  $c_k$  and  $c'_k$ . The next claim shows these constants are at most  $2^{3-k}$ .

*Claim 2.5.* For each  $k \geq 3$  there exists an  $\mathbb{N}$ -valued  $k$ -step stationary and  $\psi$ -mixing Markov chain  $\{A_n\}_{n=1}^\infty$  with  $\dim_H(\nu) \geq 1 - 2^{3-k}$ , where  $\nu$  is the distribution of  $[A_1, A_2, \dots]$ .

The main ingredient in the proof of Theorem 2.3 is Theorem 2.6 stated below, for which we need some more notations. Let  $T : X \rightarrow X$  be the Gauss map, which is defined by

$$Tx = \frac{1}{x} \pmod{1} \text{ for } x \in X.$$

Denote by  $\mu_G$  the Gauss measure, which satisfies

$$\mu_G(E) = \frac{1}{\log 2} \int_E \frac{dx}{1+x} \text{ for every Borel set } E \subset X.$$

It is well known that  $\mu_G$  is invariant and ergodic with respect to  $T$ . For  $(a_1, \dots, a_k) = \mathbf{a} \in \mathbb{N}^k$  set

$$I_{\mathbf{a}} = \{x \in X : \alpha_i(x) = a_i \text{ for each } 1 \leq i \leq k\},$$

and define  $\mathbb{I}_{\mathbf{a}} : X \rightarrow \{0, 1\}$  by

$$\mathbb{I}_{\mathbf{a}}(x) = \begin{cases} 1 & , \text{ if } x \in I_{\mathbf{a}} \\ 0 & , \text{ if } x \notin I_{\mathbf{a}} \end{cases} \text{ for } x \in X.$$

Given  $L > 1$  denote by  $\mathcal{Q}_L$  the set of maps  $q : \mathbb{N} \rightarrow \mathbb{N}$  with

$$q(n+1) > q(n) \text{ for each } n \in \mathbb{N}$$

and

$$\liminf_{n \rightarrow \infty} \frac{q(n)}{n} < L.$$

For  $q \in \mathcal{Q}_L$ ,  $\mathbf{a} \in \cup_{k=1}^\infty \mathbb{N}^k$  and  $\delta > 0$  define

$$\Gamma_{q, \mathbf{a}}^\delta = \{x \in X : \liminf_{n \rightarrow \infty} \left| \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\mathbf{a}}(T^{q(i)}x) - \mu_G(I_{\mathbf{a}}) \right| > \delta\}.$$

**Theorem 2.6.** *For every  $L > 1$  and  $\delta > 0$  there exists  $0 < c_{L, \delta} < 1$  with*

$$\sup\{\dim_H(\Gamma_{q, \mathbf{a}}^\delta) : q \in \mathcal{Q}_L, \mathbf{a} \in \cup_{k=1}^\infty \mathbb{N}^k\} \leq 1 - c_{L, \delta}.$$

*Remark 2.7.* The proof of theorem 2.6 resembles the proof of the main result (Theorem 2.1) of [KPW]. There an upper bound, which depends only on  $\delta$ , is obtained

for the dimension of sets of the form

$$(2.1) \quad \{x \in X : \limsup_{n \rightarrow \infty} \left| \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\mathbf{a}}(T^i x) - \mu_G(I_{\mathbf{a}}) \right| > \delta\}.$$

Here we need to consider the families  $\mathcal{Q}_L$ , and the more general averages

$$\frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\mathbf{a}}(T^{q(i)} x),$$

due to the lack of stationarity. As a result we must define  $\Gamma_{q, \mathbf{a}}^\delta$  with  $\liminf$ , as opposed to the sets (2.1) which are defined with  $\limsup$ .

### 3. PROOF OF THEOREM 2.6

The following large deviations estimate will be needed. Its proof is almost identical to the proof of Lemma 3.1 from [KPW], but we include it here for completeness.

**Lemma 3.1.** *Suppose  $\mathbf{S} = \{\eta_n\}_{n=1}^\infty$  is a stationary and  $*$ -mixing sequence of random variables. Let  $k \geq 1$  and  $F : \mathbb{R}^k \rightarrow \{0, 1\}$ , set*

$$p = \mathbb{P}\{F(\eta_1, \dots, \eta_k) = 1\},$$

*and let  $q : \mathbb{N} \rightarrow \mathbb{N}$  be strictly increasing. Then for every  $\delta > 0$  there exists a constant  $M = M(\mathbf{S}, \delta, k) > 1$ , independent of  $q$  and  $F$ , such that for every  $n \geq 1$ ,*

$$\mathbb{P} \left\{ \left| \frac{1}{n} \sum_{i=1}^n F(\eta_{q(i)}, \dots, \eta_{q(i)+k-1}) - p \right| > \delta \right\} \leq M \cdot e^{-n/M}.$$

*Proof.* Fix  $\delta > 0$ , then since  $\mathbf{S}$  is  $*$ -mixing there exists  $M \in \mathbb{N}$  with

$$(3.1) \quad |\mathbb{P}(E_1 \cap E_2) - \mathbb{P}(E_1)\mathbb{P}(E_2)| \leq \frac{\delta^2}{2} \mathbb{P}(E_1)\mathbb{P}(E_2)$$

for each  $m \geq 1$ ,  $E_1 \in \sigma\{\eta_1, \dots, \eta_{m+k-1}\}$  and  $E_2 \in \sigma\{\eta_{m+M}, \eta_{m+M+1}, \dots\}$ . For  $i \geq 1$  set  $\xi_i = F(\eta_i, \dots, \eta_{i+k-1})$ , fix  $n \geq M$ , and write

$$A_n = \left\{ \left| \frac{1}{n} \sum_{i=1}^n \xi_{q(i)} - p \right| > \delta \right\}.$$

Let  $N$  be the integral part of  $n/M$ , and for  $1 \leq j \leq M$  set

$$B_{n,j} = \left\{ \left| \frac{1}{N} \sum_{i=0}^{N-1} \xi_{q(j+iM)} - p \right| > \delta - \frac{1}{N} \right\}.$$

Clearly  $A_n \subset \cup_{j=1}^M B_{n,j}$ , hence

$$(3.2) \quad \mathbb{P}(A_n) \leq \sum_{j=1}^M \mathbb{P}(B_{n,j}).$$

Fix  $1 \leq j \leq M$ , and for  $\epsilon_0, \dots, \epsilon_{N-1} \in \{0, 1\}$  write

$$\mathcal{C}_{\epsilon_0, \dots, \epsilon_{N-1}} = \{\xi_{q(j+iM)} = \epsilon_i \text{ for each } 0 \leq i < N\}.$$

Let  $\zeta_0, \zeta_1, \dots$  be independent  $\{0, 1\}$ -valued random variables with mean  $p$ . Since  $q$  is strictly increasing it follows easily from (3.1) that,

$$\begin{aligned} \mathbb{P}(\mathcal{C}_{\epsilon_0, \dots, \epsilon_{N-1}}) &\leq \left(1 + \frac{\delta^2}{2}\right)^N \cdot \prod_{i=0}^{N-1} \mathbb{P}\{\xi_{q(j+iM)} = \epsilon_i\} \\ &\leq e^{\delta^2 N/2} \cdot \mathbb{P}\{\zeta_i = \epsilon_i \text{ for each } 0 \leq i < N\}. \end{aligned}$$

Set  $Z = \sum_{i=0}^{N-1} \zeta_i$ , then  $Z$  is a binomial random variable with parameters  $N$  and  $p$ , and

$$\begin{aligned} \mathbb{P}(B_{n,j}) &= \sum_{|\sum_{i=0}^{N-1} \epsilon_i - Np| > N\delta - 1} \mathbb{P}(\mathcal{C}_{\epsilon_0, \dots, \epsilon_{N-1}}) \\ (3.3) \quad &\leq e^{\delta^2 N/2} \cdot \mathbb{P}\{|Z - Np| > N\delta - 1\}. \end{aligned}$$

By the exponential estimate for the binomial distribution (see e.g. Cor. A.1.7 in [AS]) we have for  $N \geq 4/\delta$ ,

$$\mathbb{P}\{|Z - Np| > N\delta - 1\} \leq 2e^{-N\delta^2}.$$

This together with (3.3) gives,

$$\mathbb{P}(B_{n,j}) \leq 2e^{-\delta^2 N/2} \text{ for each } 1 \leq j \leq M.$$

The lemma now follows from (3.2).  $\square$

As mentioned in Remark 2.2, the sequence  $\{\alpha_i\}_{i=1}^\infty$  is  $\psi$ -mixing with respect to  $\mu_G$ . From this and Lemma 3.1 we get the following corollary.

**Corollary 3.2.** *Given  $k \geq 1$  and  $\delta > 0$  there exists a constant  $M = M(\delta, k) > 1$ , such that for every strictly increasing  $q : \mathbb{N} \rightarrow \mathbb{N}$ ,  $\mathbf{a} \in \mathbb{N}^k$  and  $n \geq 1$ ,*

$$\mu_G \left\{ x \in X : \left| \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\mathbf{a}}(T^{q(i)}x) - \mu_G(I_{\mathbf{a}}) \right| > \delta \right\} \leq M \cdot e^{-n/M}.$$

Given  $x \in X$  and  $n \geq 1$  write  $J_n(x) = I_{(\alpha_1(x), \dots, \alpha_n(x))}$ . Let  $\mathcal{L}$  be the Lebesgue measure, and write  $|I| = \mathcal{L}(I)$  for  $I \subset X$ . For  $s \geq 0$  let  $\mathcal{H}^s$  be the  $s$ -dimensional Hausdorff measure on  $X$ . For  $\eta > 0$  and  $E \subset X$  write

$$\mathcal{H}_\eta^s(E) = \inf \left\{ \sum_{i=1}^\infty |I_i|^s : E \subset \bigcup_{i=1}^\infty I_i \text{ and } |I_i| \leq \eta \right\},$$

then

$$\lim_{\eta \downarrow 0} \mathcal{H}_\eta^s(E) = \mathcal{H}^s(E).$$

Given  $n \geq 1$  write

$$\beta_n = \sup\{|I_{\mathbf{a}}| : \mathbf{a} \in \mathbb{N}^n\},$$

then  $\beta_n \xrightarrow{n} 0$ .

*Proof of Theorem 2.6.* Let  $\delta > 0$ ,  $L > 1$ ,  $q \in \mathcal{Q}_L$ ,  $k \geq 1$  and  $\mathbf{a} \in \mathbb{N}^k$ . Given  $\lambda > 0$  set,

$$\mathcal{E}_\lambda := \cap_{j=1}^\infty \cup_{n=j}^\infty \{x \in X : |J_n(x)| \leq e^{-\lambda n}\}.$$

By Theorem 4.1 in [KPW] there exists  $\lambda > 0$  with  $\dim_H \mathcal{E}_\lambda < 1$ . For  $N \geq 1$  set

$$\Gamma_{q,\mathbf{a}}^{\delta,N} := \left\{ x \in X : \begin{array}{l} |\frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\mathbf{a}}(T^{q(i)}x) - \mu_G(I_{\mathbf{a}})| > \delta, \\ |J_{q(n)+k}(x)| \geq e^{-\lambda(q(n)+k)} \end{array} \text{ for all } n \geq N \right\},$$

then

$$(3.4) \quad \Gamma_{q,\mathbf{a}}^\delta \setminus \mathcal{E}_\lambda \subset \cup_{N=1}^\infty \Gamma_{q,\mathbf{a}}^{\delta,N}.$$

Fix  $N \geq 1$  and for  $n \geq 1$  set

$$\Upsilon_{q,\mathbf{a}}^{\delta,n} := \left\{ x \in X : \begin{array}{l} |\frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\mathbf{a}}(T^{q(i)}x) - \mu_G(I_{\mathbf{a}})| > \delta, \\ |J_{q(n)+k}(x)| \geq e^{-\lambda(q(n)+k)} \end{array} \right\},$$

then  $\Gamma_{q,\mathbf{a}}^{\delta,N} \subset \Upsilon_{q,\mathbf{a}}^{\delta,n}$  for all  $n \geq N$ .

Let  $M = M(\delta, k) > 1$  be as in Corollary 3.2, set  $s = 1 - \frac{1}{\lambda LM}$  and let  $\eta > 0$ . From  $q \in \mathcal{Q}_L$  we get  $\liminf_{n \rightarrow \infty} \frac{q(n)}{n} < L$ . From this and  $\beta_n \xrightarrow{n} 0$  it follows that there exists  $n \geq N$  such that  $\beta_n < \eta$  and  $q(n) < nL$ . By the definition of  $\Upsilon_{q,\mathbf{a}}^{\delta,n}$  there exists  $B_n \subset \mathbb{N}^{q(n)+k}$  with  $\Upsilon_{q,\mathbf{a}}^{\delta,n} = \cup_{\mathbf{b} \in B_n} I_{\mathbf{b}}$ . From Corollary 3.2 we get

$$\mu_G(\Upsilon_{q,\mathbf{a}}^{\delta,n}) \leq M \cdot e^{-n/M}.$$

Since

$$r := \min_{x \in [0,1]} \frac{d\mu_G}{d\mathcal{L}}(x) > 0,$$

it follows

$$(3.5) \quad \sum_{\mathbf{b} \in B_n} |I_{\mathbf{b}}| = \mathcal{L}(\Upsilon_{q,\mathbf{a}}^{\delta,n}) \leq r^{-1} \cdot \mu_G(\Upsilon_{q,\mathbf{a}}^{\delta,n}) \leq r^{-1} M \cdot e^{-n/M}.$$

From

$$\Gamma_{q,\mathbf{a}}^{\delta,N} \subset \Upsilon_{q,\mathbf{a}}^{\delta,n} = \cup_{\mathbf{b} \in B_n} I_{\mathbf{b}}$$

and since  $|I_{\mathbf{b}}| \leq \beta_n < \eta$  for every  $\mathbf{b} \in B_n$ ,

$$(3.6) \quad \mathcal{H}_\eta^s(\Gamma_{q,\mathbf{a}}^{\delta,N}) \leq \sum_{\mathbf{b} \in B_n} |I_{\mathbf{b}}|^s \leq \left( \inf_{\mathbf{b} \in B_n} |I_{\mathbf{b}}| \right)^{s-1} \cdot \sum_{\mathbf{b} \in B_n} |I_{\mathbf{b}}|.$$

By the definition of  $\Upsilon_{q,\mathbf{a}}^{\delta,n}$ ,

$$|I_{\mathbf{b}}| \geq e^{-\lambda(q(n)+k)} \text{ for every } \mathbf{b} \in B_n.$$

Hence from (3.6), (3.5),  $q(n) < nL$  and  $s = 1 - \frac{1}{\lambda LM}$ ,

$$\begin{aligned}\mathcal{H}_\eta^s(\Gamma_{q,\mathbf{a}}^{\delta,N}) &\leq e^{\lambda(q(n)+k)(1-s)} \cdot r^{-1}M \cdot e^{-n/M} \\ &\leq r^{-1}Me^{\lambda k} \cdot \exp(n(\lambda L(1-s) - M^{-1})) = r^{-1}Me^{\lambda k}.\end{aligned}$$

As this holds for every  $\eta > 0$

$$\mathcal{H}^s(\Gamma_{q,\mathbf{a}}^{\delta,N}) = \lim_{\eta \downarrow 0} \mathcal{H}_\eta^s(\Gamma_{q,\mathbf{a}}^{\delta,N}) \leq r^{-1}Me^{\lambda k} < \infty,$$

and so

$$\dim_H(\Gamma_{q,\mathbf{a}}^{\delta,N}) \leq s = 1 - \frac{1}{\lambda LM}.$$

As this holds for every  $N \geq 1$  it follows from (3.4) that,

$$(3.7) \quad \dim_H(\Gamma_{q,\mathbf{a}}^\delta \setminus \mathcal{E}_\lambda) \leq \sup_{N \geq 1} \dim_H(\Gamma_{q,\mathbf{a}}^{\delta,N}) \leq 1 - \frac{1}{\lambda L \cdot M(\delta, k)}.$$

We shall now complete the proof of the theorem. We continue to fix  $\delta > 0$  and  $L > 1$ . Let

$$k_\delta = \inf\{k \geq 1 : \sup_{\mathbf{a} \in \mathbb{N}^k} \mu_G(I_{\mathbf{a}}) < \frac{\delta}{2}\},$$

then clearly  $k_\delta < \infty$ . For  $q \in \mathcal{Q}_L$ ,  $k \geq k_\delta$  and  $(a_1, \dots, a_k) = \mathbf{a} \in \mathbb{N}^k$ ,

$$(3.8) \quad \Gamma_{q,\mathbf{a}}^{\delta/2} \supset \left\{x \in X : \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\mathbf{a}}(T^{q(i)}x) > \delta\right\} \supset \Gamma_{q,\mathbf{a}}^\delta.$$

Set  $\mathbf{a}_\delta = (a_1, \dots, a_{k_\delta})$ , then since  $\mathbb{I}_{\mathbf{a}_\delta} \geq \mathbb{I}_{\mathbf{a}}$  it follows from (3.8) that  $\Gamma_{q,\mathbf{a}_\delta}^{\delta/2} \supset \Gamma_{q,\mathbf{a}}^\delta$ , and so

$$\dim_H(\Gamma_{q,\mathbf{a}_\delta}^{\delta/2}) \geq \dim_H(\Gamma_{q,\mathbf{a}}^\delta).$$

This together with (3.7) gives

$$\begin{aligned}\sup\{\dim_H(\Gamma_{q,\mathbf{a}}^\delta) : q \in \mathcal{Q}_L, \mathbf{a} \in \cup_{k=1}^\infty \mathbb{N}^k\} \\ \leq \max\{\dim_H(\mathcal{E}_\lambda), \max_{1 \leq k \leq k_\delta} (1 - \frac{1}{\lambda L \cdot M(\delta/2, k)})\} < 1,\end{aligned}$$

which completes the proof of the theorem.  $\square$

#### 4. PROOF OF THE MAIN RESULT

*Proof of Theorem 2.3.* Fix  $k \geq 0$ , let  $\{A_n\}_{n=1}^\infty$  an  $\mathbb{N}$ -valued  $k$ -step Markov chain which is  $*$ -mixing, and let  $\nu$  be the distribution of  $[A_1, A_2, \dots]$ . Given words  $\mathbf{a} \in \mathbb{N}^m$  and  $\mathbf{b} \in \mathbb{N}^l$  we denote by  $\mathbf{ab} \in \mathbb{N}^{m+l}$  their concatenation. As noted in observation 2.2 in [KPW], the continued fraction digits under  $\mu_G$  do not form a  $k$ -step Markov chain. It follows that there exist  $m \in \mathbb{N}$ ,  $\mathbf{a} \in \mathbb{N}^k$ ,  $\mathbf{b} \in \mathbb{N}^m$  and  $c \in \mathbb{N}$  with

$$\frac{\mu_G(I_{\mathbf{bac}})}{\mu_G(I_{\mathbf{ba}})} \neq \frac{\mu_G(I_{\mathbf{ac}})}{\mu_G(I_{\mathbf{a}})},$$



and so

$$(4.1) \quad \delta := \left| \mu_G(I_{\mathbf{bac}}) - \frac{\mu_G(I_{\mathbf{ac}}) \cdot \mu_G(I_{\mathbf{ba}})}{\mu_G(I_{\mathbf{a}})} \right| > 0.$$

If  $k = 0$ , i.e. when  $A_1, A_2, \dots$  are independent,  $\mathbf{a}$  is the empty word and  $I_{\mathbf{a}} = X$ . Let  $\mu_G(I_{\mathbf{a}}) > \epsilon > 0$  be such that if  $p_1, p_2, p_3 \in [0, 1]$  satisfy

$$|p_1 - \mu_G(I_{\mathbf{ac}})|, |p_2 - \mu_G(I_{\mathbf{ba}})|, |p_3 - \mu_G(I_{\mathbf{a}})| \leq \epsilon,$$

then

$$(4.2) \quad \left| \frac{p_1 \cdot p_2}{p_3} - \frac{\mu_G(I_{\mathbf{ac}}) \cdot \mu_G(I_{\mathbf{ba}})}{\mu_G(I_{\mathbf{a}})} \right| < \frac{\delta}{2}.$$

For each  $i \geq 1$  and  $\mathbf{d} \in \cup_{k=1}^{\infty} \mathbb{N}^k$  denote by  $E_{\mathbf{d},i}$  the event

$$\{A_i \dots A_{i+|\mathbf{d}|-1} = \mathbf{d}\},$$

where  $|\mathbf{d}|$  stands for the length of  $\mathbf{d}$ , and set  $p_{\mathbf{d},i} := \mathbb{P}(E_{\mathbf{d},i})$ . Let  $\mathbf{d} \in \cup_{k=1}^{\infty} \mathbb{N}^k$  and assume

$$\limsup_n \frac{1}{n} \# \{1 \leq i \leq n : p_{\mathbf{d},i} < \mu_G(I_{\mathbf{d}}) - \epsilon\} > \frac{1}{10},$$

then there exists  $q \in \mathcal{Q}_{10}$  with

$$(4.3) \quad p_{\mathbf{d},q(i)} < \mu_G(I_{\mathbf{d}}) - \epsilon \text{ for all } i \geq 1.$$

Since  $\{A_n\}_{n=1}^{\infty}$  is  $*$ -mixing it is evident from the definition that  $\{1_{E_{\mathbf{d},q(i)}}\}_{i=1}^{\infty}$  is also  $*$ -mixing, where  $1_E$  denotes the indicator of the event  $E$ . By the law of large numbers for sums of  $*$ -mixing bounded random variables (see Theorem 2 in [BHK]),

$$\lim_n \frac{1}{n} \sum_{i=1}^n (1_{E_{\mathbf{d},q(i)}} - p_{\mathbf{d},q(i)}) = 0 \text{ almost surely.}$$

Hence for  $\nu$ -a.e.  $x \in X$ ,

$$\lim_n \left| \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\mathbf{d}}(T^{q(i)}x) - \frac{1}{n} \sum_{i=1}^n p_{\mathbf{d},q(i)} \right| = 0.$$

From this and (4.3) we get that for  $\nu$ -a.e.  $x \in X$ ,

$$\liminf_{n \rightarrow \infty} \left| \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\mathbf{d}}(T^{q(i)}x) - \mu_G(I_{\mathbf{d}}) \right| = \liminf_{n \rightarrow \infty} \left| \frac{1}{n} \sum_{i=1}^n p_{\mathbf{d},q(i)} - \mu_G(I_{\mathbf{d}}) \right| \geq \epsilon,$$

which implies  $\nu(\Gamma_{q,\mathbf{d}}^{\epsilon/2}) = 1$ . Now by Theorem 2.6

$$\dim_H(\nu) \leq \dim_H(\Gamma_{q,\mathbf{d}}^{\epsilon/2}) \leq 1 - c_{10,\epsilon/2}.$$

In a similar manner it can be shown that  $\dim_H(\nu) \leq 1 - c_{10,\epsilon/2}$  if

$$\limsup_n \frac{1}{n} \# \{1 \leq i \leq n : p_{\mathbf{d},i} > \mu(I_{\mathbf{d}}) + \epsilon\} > \frac{1}{10}.$$

It follows that we can assume

$$\liminf_n \frac{1}{n} \# \left\{ 1 \leq i \leq n : \begin{array}{l} |p_{\mathbf{ba},i} - \mu(I_{\mathbf{ba}})| \leq \epsilon, \\ |p_{\mathbf{ac},i+m} - \mu(I_{\mathbf{ac}})| \leq \epsilon, \\ |p_{\mathbf{a},i+m} - \mu(I_{\mathbf{a}})| \leq \epsilon \end{array} \right\} > \frac{1}{10},$$

and so there exists  $q \in \mathcal{Q}_{10}$  with

$$(4.4) \quad |p_{\mathbf{ba},q(i)} - \mu_G(I_{\mathbf{ba}})|, |p_{\mathbf{ac},q(i)+m} - \mu_G(I_{\mathbf{ac}})|, |p_{\mathbf{a},q(i)+m} - \mu_G(I_{\mathbf{a}})| \leq \epsilon$$

for all  $i \geq 1$ . Since  $\{A_n\}_{n=1}^\infty$  is a Markov chain of order  $k$

$$(4.5) \quad p_{\mathbf{bac},q(i)} = \frac{p_{\mathbf{ba},q(i)} \cdot p_{\mathbf{ac},q(i)+m}}{p_{\mathbf{a},q(i)+m}} \text{ for } i \geq 1,$$

where  $p_{\mathbf{a},q(i)+m} > 0$  by (4.4) and  $\mu_G(I_{\mathbf{a}}) > \epsilon$ . The sequence  $\{1_{E_{\mathbf{bac},q(i)}}\}_{i=1}^\infty$  is  $*$ -mixing, so by the law of large numbers for sums of  $*$ -mixing random variables,

$$\lim_n \frac{1}{n} \sum_{i=1}^n (1_{E_{\mathbf{bac},q(i)}} - p_{\mathbf{bac},q(i)}) = 0 \text{ almost surely.}$$

It follows that for  $\nu$ -a.e.  $x \in X$ ,

$$\lim_n \left| \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\mathbf{bac}}(T^{q(i)}x) - \frac{1}{n} \sum_{i=1}^n p_{\mathbf{bac},q(i)} \right| = 0.$$

From this, (4.5), (4.4), (4.2) and (4.1) we get that for  $\nu$ -a.e.  $x \in X$ ,

$$\begin{aligned} & \liminf_n \left| \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\mathbf{bac}}(T^{q(i)}x) - \mu_G(I_{\mathbf{bac}}) \right| \\ &= \liminf_n \left| \frac{1}{n} \sum_{i=1}^n p_{\mathbf{bac},q(i)} - \mu_G(I_{\mathbf{bac}}) \right| \\ &\geq \left| \mu_G(I_{\mathbf{bac}}) - \frac{\mu_G(I_{\mathbf{ac}}) \cdot \mu_G(I_{\mathbf{ba}})}{\mu_G(I_{\mathbf{a}})} \right| \\ & - \limsup_n \frac{1}{n} \sum_{i=1}^n \left| \frac{p_{\mathbf{ba},q(i)} \cdot p_{\mathbf{ac},q(i)+m}}{p_{\mathbf{a},q(i)+m}} - \frac{\mu_G(I_{\mathbf{ac}}) \cdot \mu_G(I_{\mathbf{ba}})}{\mu_G(I_{\mathbf{a}})} \right| \geq \delta/2. \end{aligned}$$

Hence  $\nu(\Gamma_{q,\mathbf{bac}}^{\delta/4}) = 1$ , and so by Theorem 2.6

$$\dim_H(\nu) \leq \dim_H(\Gamma_{q,\mathbf{bac}}^{\delta/4}) \leq 1 - c_{10,\delta/4}.$$

This completes the proof of the theorem.  $\square$

## 5. CONSTRUCTION OF THE MEASURES $\nu_K$

In the proof below we use the notation for the Kolmogorov-Sinai entropy from Chapter 4 of [Wa2]. In particular the entropy of a Borel probability measure  $\theta$  on  $X$ , with respect to a countable Borel partition  $\xi$  of  $X$ , is denoted by  $H_\theta(\xi)$ . If  $\mathcal{F}$  is a sub- $\sigma$ -algebra of the Borel  $\sigma$ -algebra of  $X$ , then  $H_\theta(\xi \mid \mathcal{F})$  is the entropy of  $\theta$  with

respect to  $\xi$  conditioned on  $\mathcal{F}$ . If  $\theta$  is  $T$ -invariant the entropy of  $T$  with respect to  $\theta$  is denoted by  $h_\theta$ . If  $\theta$  is also ergodic we write  $\gamma_\theta$  for the Lyapunov exponent of the system  $(X, T, \theta)$ , i.e.

$$\gamma_\theta = \int_X \log |T'(x)| d\theta(x) = -2 \int_X \log x d\theta(x).$$

Given  $a_1, \dots, a_m \in \mathbb{N}$  we denote by  $[a_1, \dots, a_m]$  the finite continued fraction which lies in  $(0, 1)$  and has coefficients  $a_1, \dots, a_m$ , i.e.

$$[a_1, \dots, a_m] = \frac{1}{a_1 + \frac{1}{a_2 + \dots + \frac{1}{a_{m-1} + \frac{1}{a_m}}}}.$$

In order to establish the  $\psi$ -mixing property in the proof of Claim 2.5 we shall need the following proposition. It follows directly from Theorem 1 in [Br].

**Proposition 5.1.** *Let  $\{A_n\}_{n=1}^\infty$  be a stationary and mixing sequence of random variables. Assume there exists a constant  $0 < C < \infty$  with*

$$C^{-1} \leq \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(E)\mathbb{P}(F)} \leq C$$

*for all  $l \geq 1$ ,  $E \in \sigma\{A_1, \dots, A_l\}$  and  $F \in \sigma\{A_{l+1}, A_{l+2}, \dots\}$ . Then  $\{A_n\}_{n=1}^\infty$  is  $\psi$ -mixing.*

*Proof of Claim 2.5.* Fix  $k \geq 3$  and for every  $\mathbf{a} \in \mathbb{N}^k$  and  $c \in \mathbb{N}$  set

$$p_{\mathbf{a}} = \mu_G(I_{\mathbf{a}}) \text{ and } p_{\mathbf{a},c} = \frac{\mu_G(I_{\mathbf{a}c})}{\mu_G(I_{\mathbf{a}})}.$$

Then  $\sum_{c \in \mathbb{N}} p_{\mathbf{a},c} = 1$  for each  $\mathbf{a} \in \mathbb{N}^k$  and  $p = \{p_{\mathbf{a}}\}_{\mathbf{a} \in \mathbb{N}^k}$  is a probability vector. Let  $\{A_n\}_{n=1}^\infty$  be the  $k$ -step  $\mathbb{N}$ -valued Markov chain corresponding to the transition probabilities  $\{p_{\mathbf{a},c}\}_{(\mathbf{a},c) \in \mathbb{N}^{k+1}}$  and initial distribution  $\{p_{\mathbf{a}}\}_{\mathbf{a} \in \mathbb{N}^k}$ . For each  $\mathbf{b} \in \mathbb{N}^{k-1}$  and  $d \in \mathbb{N}$

$$\sum_{c \in \mathbb{N}} p_{c\mathbf{b}} \cdot p_{c\mathbf{b},d} = \sum_{c \in \mathbb{N}} \mu_G(I_{c\mathbf{b}}) \cdot \frac{\mu_G(I_{c\mathbf{b}d})}{\mu_G(I_{c\mathbf{b}})} = \mu_G(T^{-1}(I_{\mathbf{b}d})) = p_{\mathbf{b}d},$$

hence  $\{A_n\}_{n=1}^\infty$  is stationary. Considering  $\{A_n\}_{n=1}^\infty$  as a 1-step Markov chain on the state space  $\mathbb{N}^k$ , it is easy to see it is irreducible and aperiodic. From this and Theorem 8.6 in [Bi] it follows  $\{A_n\}_{n=1}^\infty$  is mixing.

Let us show  $\{A_n\}_{n=1}^\infty$  is in fact  $\psi$ -mixing. From (3.22) in chapter 3 of [EW] it follows there exists a constant  $1 < C < \infty$  with,

$$(5.1) \quad C^{-l} \leq \frac{\mu_G(I_{(a_1, \dots, a_l)})}{\mu_G(I_{a_1}) \cdot \dots \cdot \mu_G(I_{a_l})} \leq C^l \text{ for } l \geq 1 \text{ and } a_1, \dots, a_l \in \mathbb{N}.$$

For  $l, m > k$ ,  $(a_1, \dots, a_l) = \mathbf{a} \in \mathbb{N}^l$  and  $(b_1, \dots, b_m) = \mathbf{b} \in \mathbb{N}^m$  set

$$R := \frac{\mathbb{P}\{A_1 \dots A_{l+m} = \mathbf{a}\mathbf{b}\}}{\mathbb{P}\{A_1 \dots A_l = \mathbf{a}\} \mathbb{P}\{A_1 \dots A_m = \mathbf{b}\}},$$

then

$$R = \frac{1}{\mu_G(I_{(b_1, \dots, b_k)})} \cdot \prod_{j=1}^k \frac{\mu_G(I_{(a_{l-k+j}, \dots, a_l, b_1, \dots, b_j)})}{\mu_G(I_{(a_{l-k+j}, \dots, a_l, b_1, \dots, b_{j-1})})}.$$

This together with (5.1) gives

$$C^{-2k(k+1)} \leq R \leq C^{2k(k+1)}.$$

From Proposition 5.1, combined with a monotone class argument, it now follows that  $\{A_n\}_{n=1}^\infty$  is  $\psi$ -mixing.

Let  $\nu$  be the distribution of  $[A_1, A_2, \dots]$ , then  $\nu$  is  $T$ -invariant and ergodic. In order to prove the claim it remains to show that  $\dim_H \nu \geq 1 - 2^{3-k}$ . Set

$$\xi = \{I_c : c \in \mathbb{N}\},$$

then it is easy to check that

$$H_\nu(\xi) = H_{\mu_G}(\xi) < \infty$$

and

$$\sum_{c \in \mathbb{N}} \nu(I_c) \log c = \sum_{c \in \mathbb{N}} \mu_G(I_c) \log c < \infty,$$

which shows  $h_\nu, \gamma_\nu, h_{\mu_G}$  and  $\gamma_{\mu_G}$  are all finite. From this and Section 2 of [BH] it follows that

$$(5.2) \quad \dim_H \nu = \frac{h_\nu}{\gamma_\nu} \text{ and } 1 = \dim_H \mu_G = \frac{h_{\mu_G}}{\gamma_{\mu_G}}.$$

Moreover, it is well known

$$(5.3) \quad \gamma_{\mu_G} = -\frac{2}{\log 2} \int \frac{\log x}{1+x} dx = \frac{\pi^2}{6 \log 2} > 2.$$

By an argument similar to the one given in Theorem 4.27 in [Wa2],

$$h_\nu = - \sum_{\mathbf{a} \in \mathbb{N}^k} \sum_{c \in \mathbb{N}} p_{\mathbf{a}} p_{\mathbf{a},c} \log p_{\mathbf{a},c}.$$

From this and the definition of conditional entropy,

$$h_\nu = H_{\mu_G}(\vee_{j=0}^k T^{-j} \xi \mid \vee_{j=0}^{k-1} T^{-j} \sigma(\xi)).$$

Now from Theorems 4.3 and 4.14 in [Wa2],

$$(5.4) \quad \begin{aligned} h_\nu &= H_{\mu_G}(\vee_{j=0}^k T^{-j} \xi) - H_{\mu_G}(\vee_{j=1}^k T^{-j} \xi) \\ &= H_{\mu_G}(\xi \mid \vee_{j=1}^k T^{-j} \sigma(\xi)) \geq H_{\mu_G}(\xi \mid \vee_{j=1}^\infty T^{-j} \sigma(\xi)) = h_{\mu_G}. \end{aligned}$$

Assume  $k$  is even for the moment, then

$$[a_1, \dots, a_k] \leq x \leq [a_1, \dots, a_k + 1]$$

for every  $(a_1, \dots, a_k) = \mathbf{a} = \mathbb{N}^k$  and  $x \in I_{\mathbf{a}}$ . It follows that,

$$\begin{aligned}
\gamma_\nu - \gamma_{\mu_G} &= -2 \int_X \log x \, d\nu(x) + 2 \int_X \log x \, d\mu_G(x) \\
&= 2 \sum_{\mathbf{a} \in \mathbb{N}^k} \left( \int_{I_{\mathbf{a}}} \log \frac{1}{x} \, d\nu(x) + \int_{I_{\mathbf{a}}} \log x \, d\mu_G(x) \right) \\
&\leq 2 \sum_{a_1, \dots, a_k \in \mathbb{N}} \left( \int_{I_{(a_1, \dots, a_k)}} \log \frac{1}{[a_1, \dots, a_k]} \, d\nu(x) + \int_{I_{(a_1, \dots, a_k)}} \log[a_1, \dots, a_k + 1] \, d\mu_G(x) \right) \\
&= 2 \sum_{a_1, \dots, a_k \in \mathbb{N}} \mu_G(I_{(a_1, \dots, a_k)}) \cdot \log \frac{[a_1, \dots, a_k + 1]}{[a_1, \dots, a_k]}.
\end{aligned}$$

Fix  $a_1, \dots, a_k \in \mathbb{N}$ , then

$$\log \frac{[a_1, \dots, a_k + 1]}{[a_1, \dots, a_k]} \leq \frac{[a_1, \dots, a_k + 1] - [a_1, \dots, a_k]}{[a_1, \dots, a_k]}.$$

Let  $p, q \in \mathbb{N}$  be with  $\gcd(p, q) = 1$  and  $\frac{p}{q} = [a_1, \dots, a_k]$ . From inequalities (3.6), (3.7) and (3.14) in [EW] it follows that  $q, p \geq 2^{(k-2)/2}$  and

$$[a_1, \dots, a_k + 1] - [a_1, \dots, a_k] \leq q^{-2}.$$

Hence

$$\log \frac{[a_1, \dots, a_k + 1]}{[a_1, \dots, a_k]} \leq \frac{1/q^2}{p/q} = \frac{1}{pq} \leq 2^{2-k},$$

and so  $\gamma_\nu - \gamma_{\mu_G} \leq 2^{3-k}$ . By exchanging between  $\gamma_{\mu_G}$  and  $\gamma_\nu$  it can be shown that  $\gamma_{\mu_G} - \gamma_\nu \leq 2^{3-k}$ . From  $k \geq 3$  and (5.3) we get  $\gamma_\nu \geq 1$ , hence

$$(5.5) \quad 1 \leq \frac{\gamma_{\mu_G}}{\gamma_\nu} + \frac{2^{3-k}}{\gamma_\nu} \leq \frac{\gamma_{\mu_G}}{\gamma_\nu} + 2^{3-k}.$$

A similar argument shows (5.5) holds when  $k$  is odd. From (5.2), (5.4) and (5.5) we now get

$$\dim_H \nu = \frac{\gamma_{\mu_G}}{\gamma_\nu} \cdot \frac{h_\nu}{\gamma_{\mu_G}} \geq (1 - 2^{3-k}) \cdot \frac{h_{\mu_G}}{\gamma_{\mu_G}} = 1 - 2^{3-k},$$

which completes the proof of the claim.  $\square$

## 6. EXTENSION OF RESULTS FOR $f$ -EXPANSIONS

With almost no change, Theorems 2.3 and 2.6 extend to the more general setup of  $f$ -expansions, which we now define. Let  $M \in \{2, 3, \dots\} \cup \{\infty\}$ . Let  $f$  be either a strictly decreasing continuous function defined on  $[1, M+1]$  with  $f(1) = 1$  and  $f(M+1) = 0$ , or a strictly increasing continuous function defined on  $[0, M]$  with  $f(0) = 0$  and  $f(M) = 1$ . For  $x \in (0, 1)$  set  $r_0(x) = x$  and  $r_{i+1}(x) = \{f^{-1}(r_i(x))\}$  for  $i \geq 0$ , where  $\{\cdot\}$  denotes the fractional part of a number. Let  $X$  be the set of all

$x \in (0, 1)$  with  $r_i(x) \neq 0$  for every  $i \geq 0$ , then  $(0, 1) \setminus X$  is clearly countable. Write

$$\mathcal{N} = \{[y] : y \in f^{-1}(0, 1)\},$$

where  $[\cdot]$  is the integer part of a number. For  $x \in X$  and  $i \geq 1$  set

$$\alpha_i(x) = [f^{-1}(r_{i-1}(x))],$$

then  $\alpha_i(x) \in \mathcal{N}$ . We shall assume that

$$(6.1) \quad x = f(\alpha_1(x) + f(\alpha_2(x) + f(\alpha_3(x) + \dots))) \text{ for all } x \in X,$$

and call the expression on the right hand side the  $f$ -expansion of  $x$ . Regularity conditions on  $f$  were given by Rényi [R], which ensure that (6.1) is satisfied. The main example of the decreasing case is  $f(x) = 1/x$ , which leads to the continued fraction expansion, and of the increasing case is  $f(x) = x/M$ , which leads to the base- $M$  expansion. For more details on  $f$ -expansions see [R], [KP], [He] and the references therein.

We use the notation  $I_{\mathbf{a}}$  and  $\mathbb{I}_{\mathbf{a}}$ , introduced in Section 2, with  $X$  and  $\alpha_i$  as defined in this section and  $\mathbf{a} \in \cup_{k=1}^{\infty} \mathcal{N}^k$ . For  $x \in (0, 1)$  set  $Tx = f^{-1}x - [f^{-1}x]$ , then  $\alpha_i(Tx) = \alpha_{i+1}(x)$  for  $x \in X$ . We shall assume that

- (1) the restriction of  $T$  to  $f(a, a+1)$  is  $C^2$  for each  $a \in \mathcal{N}$ ;
- (2) there exists  $\ell \in \mathbb{N}$  and  $\beta > 0$  with  $|(T^\ell)'(x)| \geq \beta$  for all  $x \in X$ ;
- (3) there exists  $1 \leq Q < \infty$  with  $\left| \frac{T''(x)}{T'(y)T'(z)} \right| \leq Q$  for all  $a \in \mathcal{N}$  and  $x, y, z \in I_a$ .

Then by Theorem 22 in [Wa1], there exists an absolutely continuous  $T$ -invariant mixing probability measure  $\mu_T$  on  $X$ , such that  $0 < \frac{d\mu_T}{d\mathcal{L}} \in C[0, 1]$ . Here, as above,  $\mathcal{L}$  is the Lebesgue measure.

For  $q \in \mathcal{Q}_L$  with  $\mathcal{Q}_L$  defined in Section 2,  $\mathbf{a} \in \cup_{k=1}^{\infty} \mathcal{N}^k$  and  $\delta > 0$  let

$$\Gamma_{q, \mathbf{a}}^\delta = \{x \in X : \liminf_{n \rightarrow \infty} \left| \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\mathbf{a}}(T^{q(i)}x) - \mu_T(I_{\mathbf{a}}) \right| > \delta\}.$$

The following theorem is an analogue of Theorem 2.6, and can be proven in exactly the same manner.

**Theorem 6.1.** *Suppose that  $T$  satisfies conditions (1)-(3) and assume, in addition, that for some  $t < 1$ ,*

$$(6.2) \quad \sup_{x \in X} \sum_{y: Ty=x} |T'(y)|^{-t} < \infty.$$

*Then for every  $L > 1$  and  $\delta > 0$  there exists  $0 < c_{f,L,\delta} < 1$  with*

$$\sup\{\dim_H(\Gamma_{q, \mathbf{a}}^\delta) : q \in \mathcal{Q}_L, \mathbf{a} \in \cup_{k=1}^{\infty} \mathcal{N}^k\} \leq 1 - c_{f,L,\delta}.$$

*Remark 6.2.* The condition (6.2) is needed in order to apply Theorem 4.1 from [KPW], as we did at the beginning of the proof of Theorem 2.6. Since  $\{\alpha_i\}_{i=1}^{\infty}$

is a  $\psi$ -mixing sequence with respect to  $\mu_T$  (see [Ad] or [He]), the large deviations estimate from Corollary 3.2 is valid for  $\mu_T$ . Now the proof of Theorem 6.1 follows almost verbatim the proof of Theorem 2.6.

An important ingredient in the proof of Theorem 2.3 is the fact that, for any  $k \geq 0$ , the continued fraction digits under  $\mu_G$  do not form a  $k$ -step Markov chain. Hence, in order to generalize Theorem 2.3 to the case of  $f$ -expansions we shall need the following lemma. For  $t \in [0, 1]$  set  $F(t) = \mu_T([0, t])$ , and let  $S = F \circ T \circ F^{-1}$ . Since  $F' = \frac{d\mu_T}{d\mathcal{L}} \in C[0, 1]$  with  $\frac{d\mu_T}{d\mathcal{L}} > 0$ ,  $F$  is a diffeomorphism of  $[0, 1]$  onto itself. Given  $a \in \mathcal{N}$  write  $\tilde{I}_a := f(a, a + 1)$ .

**Lemma 6.3.** *Assume the digits  $\{\alpha_i\}_{i=1}^\infty$  of the  $f$ -expansion are not independent under  $\mu_T$ . Then  $\{\alpha_i\}_{i=1}^\infty$  do not form a  $k$ -step Markov chain under  $\mu_T$  for any  $k \geq 1$ .*

*Proof.* Note that  $F\mu_T = \mathcal{L}$  and  $S\mathcal{L} = \mathcal{L}$ . From the chain rule it follows that for every  $a \in \mathcal{N}$  and  $x \in F\tilde{I}_a$ ,

$$S'(x) = F'(TF^{-1}x)T'(F^{-1}x)(F'(F^{-1}x))^{-1},$$

and so  $S'$  is continuous on  $F\tilde{I}_a$ . Let  $\beta_1 : FX \rightarrow \mathcal{N}$  be such that  $\beta_1(x) = a$  for  $a \in \mathcal{N}$  and  $x \in FI_a$ . For  $i \geq 1$  set  $\beta_i = \beta_1 \circ S^{i-1}$ , then  $\beta_i = \alpha_i \circ F^{-1}$ . Given  $(a_1, \dots, a_l) = \mathbf{a} \in \mathcal{N}^l$  let

$$J_{\mathbf{a}} = \{x \in FX : \beta_i(x) = a_i \text{ for } 1 \leq i \leq l\},$$

then  $J_{\mathbf{a}} = FI_{\mathbf{a}}$ . Note that

$$(6.3) \quad \mathcal{L}(J_{\mathbf{a}}) = \mu_T(I_{\mathbf{a}}) \text{ for every } l \geq 1 \text{ and } \mathbf{a} \in \mathcal{N}^l.$$

Let  $k \geq 1$  and assume by contradiction that  $\{\alpha_i\}_{i=1}^\infty$  forms a  $k$ -step Markov chain under  $\mu_T$ . From this, since  $\{\alpha_i\}_{i=1}^\infty$  are not independent, and from (6.3), it follows that under  $\mathcal{L}$  the variables  $\{\beta_i\}_{i=1}^\infty$  form a stationary  $k$ -step Markov chain but are not independent. Since  $\{\beta_i\}_{i=1}^\infty$  is a stationary  $k$ -step Markov chain,

$$\mathcal{L}\{\beta_1 = c \mid S^{-1}(J_{\mathbf{b}})\} = \mathcal{L}\{\beta_1 = c \mid S^{-1}(J_{\mathbf{bf}})\}$$

for every  $c \in \mathcal{N}$ ,  $\mathbf{b} \in \mathcal{N}^k$ ,  $l \geq 1$  and  $\mathbf{f} \in \mathcal{N}^l$ . It follows there exist  $c \in \mathcal{N}$  and  $\mathbf{b}, \mathbf{d} \in \mathcal{N}^k$  with,

$$(6.4) \quad \mathcal{L}\{\beta_1 = c \mid S^{-1}(J_{\mathbf{b}})\} \neq \mathcal{L}\{\beta_1 = c \mid S^{-1}(J_{\mathbf{d}})\},$$

otherwise it would hold that  $\{\beta_i\}_{i=1}^\infty$  are independent under  $\mathcal{L}$ .

It is not hard to see that for  $\mathcal{L}$ -a.e.  $x \in J_c$ ,

$$(6.5) \quad \mathcal{L}\{\beta_1 = c \mid \sigma\{\beta_2, \beta_3, \dots\}\}(x) = (S'(x))^{-1},$$

where the left hand side is the conditional  $\mathcal{L}$ -probability of the event  $\{\beta_1 = c\}$  with respect to the  $\sigma$ -algebra  $\sigma\{\beta_2, \beta_3, \dots\}$ . Let  $\mathbf{a} \in \mathcal{N}^k$ . Then since  $\{\beta_i\}_{i=1}^\infty$  is a  $k$ -step Markov chain under  $\mathcal{L}$ , it follows for  $\mathcal{L}$ -a.e.  $x \in J_{c\mathbf{a}}$  that

$$\begin{aligned} & \mathcal{L}\{\beta_1 = c \mid \sigma\{\beta_2, \beta_3, \dots\}\}(x) \\ &= \mathcal{L}\{\beta_1 = c \mid \sigma\{\beta_2, \dots, \beta_{k+1}\}\}(x) = \mathcal{L}\{\beta_1 = c \mid S^{-1}(J_{\mathbf{a}})\}. \end{aligned}$$

This together with (6.5) shows that

$$(6.6) \quad (S'(x))^{-1} = \mathcal{L}\{\beta_1 = c \mid S^{-1}(J_{\mathbf{a}})\} \text{ for } \mathcal{L}\text{-a.e. } x \in J_{c\mathbf{a}}.$$

Since  $S'$  is continuous on  $F\tilde{I}_c$  and

$$F(\tilde{I}_c \cap X) = \cup_{\mathbf{a} \in \mathcal{N}^k} J_{c\mathbf{a}},$$

it follows easily from (6.6) that  $S'$  must be constant on  $F\tilde{I}_c$ . On the other hand, by (6.4) and (6.6) this is not possible. We have thus reached a contradiction, which shows that  $\{\alpha_i\}_{i=1}^\infty$  does not form a  $k$ -step Markov chain under  $\mu_T$ .  $\square$

*Remark 6.4.* In Proposition 7.1 from [KPW] it is shown that  $\{\alpha_i\}_{i=1}^\infty$  are independent under  $\mu_T$  if and only if  $S$  is linear on  $F\tilde{I}_a$  for each  $a \in \mathcal{N}$ . From this and Lemma 6.3 it follows that if  $S$  is not linear on  $F\tilde{I}_a$  for some  $a \in \mathcal{N}$ , then  $\{\alpha_i\}_{i=1}^\infty$  do not form a  $k$ -step Markov chain under  $\mu_T$  for any  $k \geq 0$ .

The following theorem is an analogue, for the case of  $f$ -expansions, of Theorem 2.3 above and Corollary 2.3 from [KPW]. It can be derived from Theorem 6.1, Theorem 2.1 in [KPW], and Lemma 6.3, by an argument similar to the one given in the proof of Theorem 2.3. Given  $a_1, a_2, \dots \in \mathcal{N}$  denote by  $[a_1, a_2, \dots]$  the unique  $x \in X$  with  $\alpha_i(x) = a_i$  for  $i \geq 1$ .

**Theorem 6.5.** *Suppose that  $T$  satisfies the conditions (1)-(3) and, in addition, that (6.2) holds for some  $t < 1$ . Assume the digits  $\{\alpha_i\}_{i=1}^\infty$  of the  $f$ -expansion are not independent under  $\mu_T$ . Let  $k \geq 0$  and let  $\{A_n\}_{n=1}^\infty$  be an  $\mathcal{N}$ -valued  $k$ -step Markov chain (when  $k = 0$  this means  $A_1, A_2, \dots$  are independent). Assume  $\{A_n\}_{n=1}^\infty$  is  $*$ -mixing or that it is stationary and ergodic. Let  $\nu$  be the distribution of the random variable  $[A_1, A_2, \dots]$ . Then  $\dim_H(\nu) \leq 1 - c_{f,k}$ , where  $0 < c_{f,k} < 1$  is a constant depending only on  $f$  and  $k$ .*

## REFERENCES

- [Ad] R. Adler, Continued fractions and Bernoulli trials, in: Ergodic Theory (A Seminar), J. Moser, E. Phillips and S. Varadhan, eds., Courant Inst. of Math. Sci. (Lect. Notes 110), 1975, New York.
- [AS] N. Alon, J. Spencer, The Probabilistic Method, Wiley, New York, 1992.
- [Bi] P. Billingsley, Probability and Measure, 3rd edn., Wiley, New York 1995.



- [BH] P.Billingsley and I.Henningsen, Hausdorff dimension of some continued-fraction sets. *Z. Wahrsch. verw. Geb.* 31 (1975), 163–173.
- [BHK] J.R.Blum, D.L.Hanson and L.H.Koopmans, On the Strong Law of Large Numbers for a Class of Stochastic Processes. *Z. Wahrsch. verw. Geb.* 2 (1963), 1-11.
- [Br] R.C.Bradley, On the  $\psi$ -mixing condition for stationary random sequences. *Trans. Amer. Math. Soc.* 276 (1983) 55-66. MR684493.
- [Ch] S.D.Chatterji, Masse, die von regelmässigen Kettenbrüchen induziert sind, *Math. Ann.*, vol. 164 (1966), pp. 113-117. MR0193079 (33:1300).
- [EW] M.Einsiedler and T.Ward, *Ergodic Theory with a View Towards Number Theory* (Graduate Texts in Mathematics, 259). Springer, London, 2011.
- [He] L.Heinrich, Mixing properties and central limit theorem for a class of non-identical piecewise monotonic  $C^2$ -transformations. *Mathematische Nachricht.* 181, 185–214 (1996).
- [KPW] Y.Kifer, Y.Peres and B.Weiss, A dimension gap for continued fractions with independent digits. *Israel J. Math.* 124(1), (2001), 61–76.
- [KP] J.R.Kinney and T.S.Pitcher, The Dimension of Some Sets Defined in Terms of  $f$ -expansions. *Z. Wahrsch. verw. Geb.* 4 (1966), 293–315
- [R] A.Rényi, Representations for real numbers and their ergodic properties, *Acta Math. Acad. Sci. Hung.*, 8 (1957), pp. 477–493.
- [Wa1] P.Walters, Invariant measures and equilibrium states for some mappings which expand distances. *Trans. Amer. Math. Soc.* 236 (1975), 121-153.
- [Wa2] P.Walters, *An introduction to ergodic theory.* Springer, Verlag, 2001.